Communication

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 43, Issue of October 23, pp. 27761–27764, 1998
Printed in U.S.A.

Oncoprotein TLS Interacts with Serine-Arginine Proteins Involved in RNA Splicing*

(Received for publication, August 18, 1998)

Liu Yang‡§, Lisa J. Embree‡§, Schickwann Tsai¶, and Dennis D. Hickstein‡§‖**

From the ‡Medical Research Service, Veterans Affairs Puget Sound Health Care System, Seattle, Washington 98108, ¶Fred Hutchinson Cancer Research Center, Seattle, Washington 98104, and §Division of Oncology and ‖Molecular Medicine, University of Washington School of Medicine, Seattle, Washington 98195

The gene encoding the human TLS protein, also termed FUS, is located at the site of chromosomal translocations in human leukemias and sarcomas where it forms a chimeric fusion gene with one of several different genes. To identify interacting partners of TLS, we screened a yeast two-hybrid cDNA library constructed from mouse hematopoietic cells using the C-terminal region of TLS in the bait plasmid. Two cDNAs encoding members of the serine-arginine (SR) family of proteins were isolated. The first SR protein is the mouse homolog of human splicing factor SC35, and the second SR member is a novel 183-amino acid protein that we term TASR (TLS-associated serine-arginine protein). cDNA cloning of human TASR indicated that mouse and human TASR have identical amino acid sequences. The interactions between TLS and these two SR proteins were confirmed by co-transfection and immunoprecipitation studies. *In vivo* splicing assays indicated that SC35 and TASR influence splice site selection of adenovirus E1A pre-mRNA. TLS may recruit SR splicing factors to specific target genes through interaction with its C-terminal region, and chromosomal translocations that truncate the C-terminal region of TLS may prevent this interaction. Thus TLS translocations may alter RNA processing and play a role in malignant transformation.

Chromosomal translocations are found frequently in leukemias as well as in malignancies of non-hematopoietic tissues. These translocations usually give rise to novel fusion genes and novel fusion proteins. To understand the role that these fusion proteins play in cellular transformation, knowledge of the function of the wild-type protein involved in the translocations is required.

We have focused these studies on understanding the function of the *TLS* gene. TLS (translocated in liposarcoma), also called FUS, was originally identified through its fusion to the *CHOP* gene (a member of the C/EBP family of transcription factors) in human myxoid liposarcoma with the t(12;16) chromosomal translocation (1, 2). In human acute myeloid leukemia with the t(16;21) translocation, the *TLS* gene is fused with the *ERG* gene (a member of the ETS family of transcription factors) (3). In translocations involving TLS, the N-terminal part of TLS is retained in the fusion protein and the C-terminal region is deleted.

TLS is a member of a closely related family of genes, including the *EWS* gene, which was originally identified in Ewing's sarcoma (4). Both TLS and EWS are involved in various types of cancers through chromosomal translocation either with other genes of the ETS family or with other transcription factors (see Ref. 5 for a review). The evidence that transcriptional activation plays a role in transformation mediated by TLS (or EWS) fusion proteins stems from the observations that the N terminus of the TLS protein is rich in glutamine, serine, and tyrosine and is a potent transactivator when fused with various transcription factors (6).

Two observations suggest that the transactivational activity of the TLS fusion protein may not be sufficient to explain the role of the fusion protein in transformation. First, a correlation between transactivation and transformation by these fusion proteins has not been demonstrated (7). Second, the C terminus of TLS contains several motifs that are suggestive of other potential TLS functions. The TLS C-terminal region contains an evolutionarily conserved RNP[1] consensus sequence and Arg-Gly-Gly (RGG) repeats, both of which have been implicated in RNA binding (8). In previous studies, TLS has been shown to bind to RNA in both the nucleus and the cytoplasm and was hypothesized to be a heterogeneous ribonuclear protein-like chaperone of RNA (9). In a recent study, TLS was found to interact with PU.1, an ETS protein capable of regulating transcription and RNA splicing. Overexpression of TLS in IW1–32 erythroid cells was shown to promote the use of the distal 5′-splice site during E1A pre-mRNA splicing. This site preference is counterpoised by PU.1, suggesting that TLS may be part of a protein network involved in the regulation of RNA processing (10).

Because the C terminus of TLS is replaced in fusion proteins as the result of chromosomal translocation, loss of TLS interactions with cellular regulators may play an important role in transformation. To identify protein molecules that interact with TLS, we used the C-terminal part of TLS protein as the bait to screen a yeast two-hybrid library derived from mouse multipotential hematopoietic cells (11). We isolated two cDNAs that encode TLS-interacting proteins with serine-arginine (SR)-rich domains. The first cDNA corresponds to the mouse homolog of the human splicing factor SC35, and the second cDNA encodes a novel protein that we term TASR (TLS-associated serine-arginine protein). The SR family of proteins has been shown previously to be involved in constitutive and regulated RNA splicing.

* This work was supported by grants from the Department of Veterans Affairs and the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF077858, AF042383, and AF047448.

** To whom correspondence should be addressed: VA Puget Sound Health Care System, 1660 S. Columbian Way, GMR 151, Seattle, WA 98108. Tel.: 206-764-2705; Fax: 206-764-2827; E-mail: dennishi@u.washington.edu.

[1] The abbreviations used are: RNP, ribonucleoprotein; SR, serine-arginine; PCR, polymerase chain reaction; EML, erythroid, myeloid, and lymphoid; RT, reverse transcription; PAGE, polyacrylamide gel electrophoresis; DOTAP, *N*-[dioleoyloxy]propyl-trimethylammonium methylsulfate.

EXPERIMENTAL PROCEDURES

Plasmid and Library Constructions—To generate a TLS bait plasmid for yeast two-hybrid screening, human TLS cDNA was PCR-amplified, digested with *Eco*RI, and cloned into the *Eco*RI-*Sma*I sites of the pBTM116 vector (12) to generate the bait plasmid pBTM-TLS in which the C-terminal region of TLS protein is fused to LexA DNA-binding protein (see Fig. 1). For expression of Flag-tagged TLS protein in COS-7 cells, the entire coding region of the TLS cDNA was subcloned in-frame into the *Eco*RI-*Sma*I sites of pSG5-Flag vector (CLONTECH) to generate pSG5-Flag-TLS. Flag-tagged luciferase expression plasmid pSG5-Flag-Luc was constructed by ligation of the pSG5-Flag vector with the luciferase insert from pGL3 plasmid (Promega). Myc-tagged expression constructs pCS2-MT-SC35 and pCS2-MT-TASR were generated by in-frame ligation of mouse SC35 and TASR coding sequences into the *Eco*RI-*Stu*I sites of pCS2-MT vector. For overexpression of the SC35 and TASR proteins in HeLa cells, full-length mouse SC35 and TASR cDNAs were cloned into the *Eco*RI-*Eco*RV sites of pCR3 vector (Invitrogen) to generate plasmids pCR3-SC35 and pCR3-TASR. Plasmid pCS3-MT-E1A was from Dr. Moreau-Gachelin and described previously (10). The yeast two-hybrid library was constructed with mRNAs from the murine hematopoietic cells with erythroid, myeloid, and lymphoid (EML) potential (11, 12).

Two-hybrid Screen—The yeast two-hybrid screen was performed as described previously (11).

cDNA Cloning—To obtain full-length mouse SC35 and TASR cDNA sequences, the SC35 and TASR inserts obtained using the yeast two-hybrid screen were used as the probe in the hybridization screening of a Uni-ZAP phage cDNA library derived from EML cells. pBluescript phagemids containing full-length SC35 and TASR cDNAs were prepared after *in vivo* excision from the Uni-ZAP XR vector and were used in sequencing reactions with Dye terminators (Applied Biosystems). To obtain human TASR cDNA, total RNA was isolated from K562 leukemia cells and used as a template for amplification by RT-PCR.

Immunoprecipitation and Western Blot Analysis—For expression of Flag- or Myc-tagged proteins, 10 μg of pSG5-Flag-expression construct and 10 μg of pCS2-Myc-expression construct were introduced into 3 × 10^6 COS-7 cells by electroporation. 48 h after electroporation, the cells were lysed with 0.6 ml of lysis buffer A (10 mM Tris·HCl, pH 7.4, 2.5 mM $MgCl_2$, 100 mM NaCl, 0.5% Triton X-100). 3 μl of polyclonal rabbit anti-Flag D8 antibody (Santa Cruz Biotechnology) or 3 μl of monoclonal mouse anti-Myc 9E10 antibody (Sigma) was first incubated with 30 μl of protein A/G agarose (Santa Cruz Biotechnology) for 50 min at 4 °C in 0.3 ml of buffer A, and the antibody-protein A/G-agarose complex was then incubated with 0.2 ml of lysate for 20 min at 4 °C with gentle rocking. After washing with radioimmune precipitation buffer 4 times, 50 μl of SDS-PAGE sample buffer was added to the agarose beads. The samples were heated at 100 °C for 3 min, 20 μl of the sample was separated by SDS-PAGE in a 10% gel, and the proteins were detected with a monoclonal mouse anti-Flag M2 antibody (Sigma) or a monoclonal mouse anti-Myc 9E10 antibody as described under "Results and Discussion." Protein bands were visualized using the ECL Western blotting analysis system (Amersham Pharmacia Biotech).

In Vivo Splicing Assay—For transient transfection of HeLa cells, 1.7 μg of pCS3-MT-E1A and 1.7 μg of pCR3-construct plus 1.7 μg of pSG5-Flag-construct were mixed with 30 μl of DOTAP (Boehringer Mannheim), and the DNA-DOTAP mixture was added to a 60-mm dish of 75% confluent HeLa cells according to the manufacturer's instructions. 40 h after addition of the DNA-DOTAP mixture, total RNA was purified from transfected HeLa cells with an RNeasy Mini Kit (Qiagen). RT-PCR amplification of various E1A isoforms was carried out as described previously (13) with 5′GAGCTTGGGCGACCTCA3′ (RR67) as the forward primer and 5′AATACGACTCACTATAG3′ (T7) as the reverse primer.

RESULTS AND DISCUSSION

The structural features of TLS, fusion proteins TLS/CHOP and TLS/ERG, and the C-terminal region of TLS used in the bait plasmid are shown (Fig. 1). TLS sequence features implicated in RNA binding include the Arg-Gly-Gly repeats (RGG) and the evolutionarily conserved ribonucleoprotein consensus sequence (RNP-CS).

In the initial yeast two-hybrid screen of 1 × 10^7 transformants, we identified two independent clones that were positive for interaction with the TLS bait. Sequencing of the first clone indicated that it was the mouse homolog of human SC35, one of the SR-rich mammalian splicing factors (14). Sequencing of the

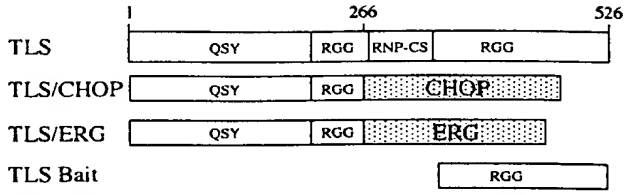


FIG. 1. **Diagrams of TLS and TLS fusion proteins.** Normal TLS protein is shown with distinct sequence features. *QSY*, glutamine-, serine-, and tyrosine-rich domain; *RGG*, regions with multiple Arg-Gly-Gly repeats; *RNP-CS*, ribonucleoprotein consensus sequence. Sites of TLS/CHOP and TLS/ERG fusions are indicated. The TLS region used as a bait in the yeast two-hybrid screen is indicated.

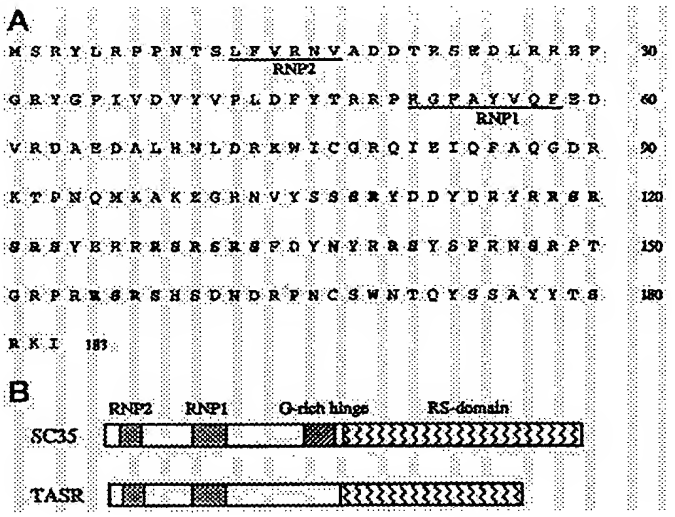


FIG. 2. **Amino acid sequence of TASR and its structural comparison to SC35.** *A*, the predicted amino acid sequence of TASR protein. Both mouse TASR (GenBank accession no. AF042383) and human TASR (GenBank accession no. AF047448) are identical at the amino acid level. RNP2 and RNP1 consensus sequences are *underlined*. Arg-Ser or Ser-Arg dipeptide repeats are shown in *boldface*. *B*, schematic comparison of TASR to SC35. The RNP consensus sequences are shown in *gray boxes*. The glycine-rich hinge is in a *hatched box*, and the RS domains are in *boxes* with *wavy lines*.

second positive clone revealed the presence of C-terminal serine-arginine repeats, whereas the remainder of the sequence did not align with sequences in the GenBank data base. In-frame translation of the cDNA insert indicated that it represented a novel TLS-associated protein containing multiple serine-arginine repeats; therefore the protein was named TASR (TLS-associated protein with serine-arginine repeats).

To obtain full-length mouse SC35 and TASR cDNAs, the corresponding cDNA inserts were used to screen an EML cDNA phage library. Several clones containing full-length SC35 and TASR cDNAs were isolated. Nucleotide sequence analysis revealed that mouse SC35 cDNA is approximately 1.9 kilobases in length (GenBank accession no. AF077858) and encodes a protein nearly identical to human SC35. Mouse TASR cDNA is approximately 3.4 kilobases in length (GenBank accession no. AF042383). The sequence surrounding the first ATG codon of mouse TASR cDNA is an excellent match to the Kozak translation initiation consensus (15). The deduced amino acid sequence of the TASR open reading frame corresponds to a protein of 183 amino acids with a calculated molecular mass of 22 kDa (Fig. 2*A*). Prominent in the TASR sequence are RNP2 and RNP1 motifs, which are signatures of RNA-binding proteins, as well as multiple serine-arginine repeats, which are a characteristic feature of RNA splicing factors. Even though a search of

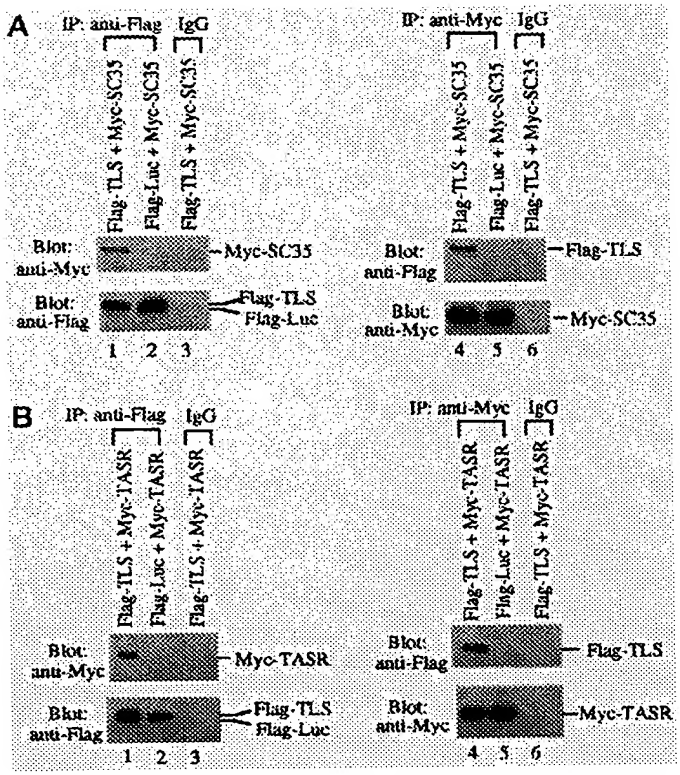


FIG. 3. **Association of SC35 and TASR with TLS in COS-7 cells.** *A*, Plasmid expressing Myc-tagged SC35 protein was co-transfected into COS-7 cells with plasmid expressing Flag-tagged TLS or Flag-tagged luciferase. The cell lysates were immunoprecipitated with a polyclonal rabbit anti-Flag antibody (*lanes 1* and *2*) or with a normal rabbit IgG (*lane 3*). The immunoprecipitates (*IP*) were isolated, subjected to gel electrophoresis, and immunoblotted with a monoclonal mouse anti-Myc antibody or a mouse monoclonal anti-Flag antibody. The reciprocal immunoprecipitation was carried out with a monoclonal mouse anti-Myc antibody (*lanes 4* and *5*) or a normal mouse IgG (*lane 6*). The precipitated proteins were separated and immunoblotted with a monoclonal mouse anti-Flag or a monoclonal mouse anti-Myc antibody. *B*, experiments as in *A* were carried out using Myc-tagged TASR.

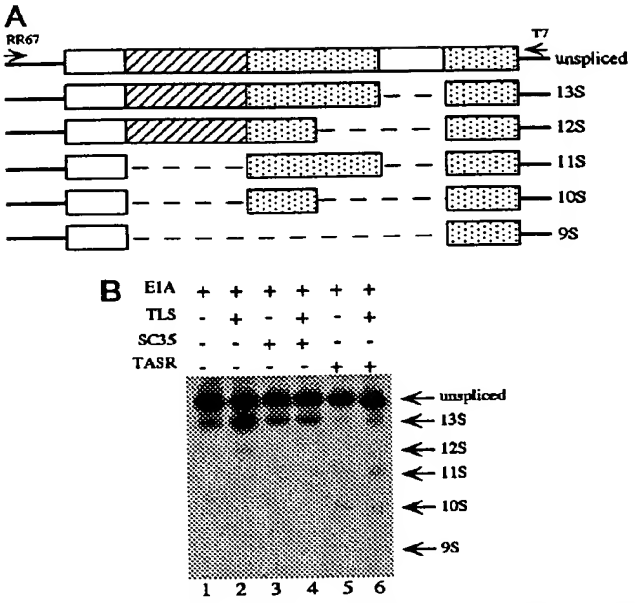


FIG. 4. **Effects of TLS, SC35, and TASR overexpression on E1A pre-mRNA splicing in HeLa cells.** *A*, diagrams of E1A pre-mRNA splicing. Various E1A splicing isoforms are shown with *dashed lines* representing spliced sequences. Primers for RT-PCR analysis are indicated by *arrows*. *B*, *in vivo* alternative splicing of E1A pre-mRNA in HeLa cells. DNA combinations for all samples are indicated at the *top*. The total amount of DNA in each transfection was kept constant through addition of control plasmids. Different E1A splicing isoforms were amplified by RT-PCR, and all E1A bands were confirmed by hybridization to a ^{32}P-labeled E1A DNA probe.

the GenBank protein data base with the BLASTp program (16) revealed that the TASR protein shares similar structural motifs with the SR family of splicing factors such as SC35, SF2/ASF (17, 18), and SRp20 (19), the conserved glycine-rich hinge in SC35 and SF2/ASF is absent in TASR protein (Fig. 2*B*).

Using RT-PCR, we have also cloned the human TASR cDNA (GenBank accession no. AF047448) from K562 leukemia cells. Sequencing analysis revealed that both mouse and human TASR coding regions are 93% identical at the nucleotide level and 100% identical at the amino acid level, indicating that TASR protein is evolutionarily conserved.

The finding that SC35 interacts with TLS in the yeast two-hybrid system suggested that these two proteins may also associate with each other in mammalian cells. To confirm their intracellular association, plasmids expressing Myc-SC35 and Flag-TLS or Flag-luciferase were co-transfected into COS-7 cells, and lysates from the co-transfected cells were used for immunoprecipitation. A polyclonal rabbit anti-Flag antibody co-immunoprecipitated Myc-SC35 and Flag-TLS (Fig. 3*A*, *lane 1*) but did not co-immunoprecipitate Myc-SC35 along with Flag-luciferase (Fig. 3*A*, *lane 2*). A normal rabbit IgG did not co-immunoprecipitate Myc-SC35 and Flag-TLS (Fig. 3*A*, *lane 3*), indicating the specificity of the protein-antibody interaction. In the reciprocal immunoprecipitation, only Flag-TLS was co-immunoprecipitated with Myc-SC35 by a monoclonal mouse anti-Myc antibody (Fig. 3*A*, *lane 4*). Flag-luciferase failed to associate with Myc-SC35 and could not be detected in the immunoprecipitates (Fig. 3*A*, *lane 5*). When a normal mouse IgG was used in the experiment, neither Flag-TLS nor Myc-SC35 was present in the immunoprecipitates (Fig. 3*A*, *lane 6*).

To confirm the intracellular association between TASR and TLS, plasmids expressing Myc-TASR and Flag-TLS or Flag-luciferase were also transfected into COS-7 cells for co-immunoprecipitation studies. Myc-TASR and Flag-TLS formed a protein complex that is recognized by specific antibodies directed against the epitope tags (Fig. 3*B*, *lanes 1* and *4*). Myc-TASR did not bind to Flag-luciferase (Fig. 3*B*, *lanes 2* and *5*), indicating the specificity of the interaction between TLS and TASR. Normal IgG did not immunoprecipitate Myc-TASR or Flag-TLS (Fig. 3*B*, *lanes 3* and *6*).

To investigate the functions of TLS, SC35, and TASR as splicing factors, we analyzed the effects of their overexpression on the alternative splicing of the adenovirus E1A pre-mRNA in HeLa cells. Differential splicing of E1A pre-mRNA generates five major isoforms (13, 12, 11, 10, and 9 S) (Fig. 4*A*) (20). When overexpressed in HeLa cells, TLS promoted splice site selection generating the 13 and 12 S isoforms (Fig. 4*B*, compare *lanes 1* and *2*). In agreement with previous reports, expression of SC35 did not increase 13 S isoform but decreased the 12 S isoform (Fig. 4*B*, compare *lanes 1* and *3*) (21). When both TLS and SC35 were co-expressed, TLS was unable to promote the use of splice sites leading to 13 and 12 S isoforms (Fig. 4*B*, compare *lanes 2* and *4*). The TASR protein also functioned as a splicing factor. In contrast to SC35, however, TASR expression decreased the 13 S isoform and promoted the use of distal 5′-splice sites leading to the 11, 10, and 9 S isoforms (Fig. 4*B*, compare *lanes 1* and *5*). When TLS and TASR were co-expressed, again the distal 5′-splice sites resulting from TASR selection were favored (Fig. 4*B*, compare *lanes 2* and *6*). Together, these results

indicate that SC35 and TASR display opposing effects on E1A pre-mRNA splicing and that the influence of TLS on E1A pre-mRNA splicing is abrogated in the presence of overexpressed SC35 or TASR.

Considerable evidence now implicates TLS in binding to RNA. First, TLS protein contains structural motifs such as RGG and RNP-CS, which are conserved in many RNA-binding proteins (8). Second, bacterially expressed full-length TLS and its C terminus both bind to *in vivo* labeled RNA (1). The fact that this binding is enriched in the poly(A)$^+$ fraction of RNA suggests that the targets for TLS are in the mRNA fraction. Third, *in vivo* UV cross-linking studies indicate that TLS binds to RNA species that turn over rapidly inside cells (9). Fourth, immunocytochemical analysis localizes TLS protein to nuclear regions outside of the nucleolus, further supporting the notion that TLS binds to non-ribosomal RNA transcripts (1, 22).

The evidence linking TLS to splicing of nascent pre-mRNA is also accumulating. TLS has been found to be associated with proteins, such as human nuclear RNP A1 and SF1, that are implicated in splicing (23, 24). Also, as mentioned in the introduction, TLS induces the preferential use of the most distal 5′-splice site during E1A pre-mRNA splicing in IW1–32 erythroid cells (10). At the present time, the mechanism by which TLS influences pre-mRNA splicing remains unclear. This study, for the first time, links TLS to the SR family of proteins, which have been extensively investigated for their role in RNA splicing.

RNA processing is clearly a primary function of SR proteins. The N-terminal RNP consensus sequence of SR proteins is required for binding to RNA, whereas the C-terminal arginine-serine-rich domain is important in mediating association with other cellular proteins (25). In addition, phosphorylation of SR proteins has also been reported to affect both protein-protein and protein-RNA interactions, thus exerting further control over the enzymatic activity and intracellular localization of these splicing factors (26, 27). Although different SR proteins may have overlapping functions in constitutive splicing, they differ from each other in their ability to regulate alternative splicing as shown by previous studies (21, 28) and by the *in vivo* splicing assays of this paper.

Important regulatory roles for SR proteins have been demonstrated by their functions in *Drosophila* development and sex determination (29–31), their involvement in T-cell activation (32), their close association with cell cycle control (33), and their interaction with key cellular enzymes such as DNA topoisomerase I (34). SR proteins have been found to be associated with various transcription units (35), linking SR proteins to the tightly coupled processes of transcription and splicing in mammalian cells (36). The association of the SR family of splicing factors with RNA polymerase II has also been described by several groups (37–39), suggesting interactions between SR proteins and the C-terminal repeat domain of RNA polymerase II.

Because of the potential functions of TLS and SR proteins, interactions between the SR family of proteins and TLS-related proteins may be an integral part of the control of normal cell growth and differentiation. Because our yeast two-hybrid screen suggests that the C-terminal region of TLS is involved in binding to SR proteins, interactions between TLS and SR proteins are likely to be altered or disrupted when this part of TLS protein is replaced through chromosomal translocations. The role that the TLS fusion proteins play in the TLS-SR protein interaction and their effect on RNA processing of TLS target genes are important questions for further study. Alterations of RNA splicing for several genes are associated with various tumor types (40). This study thus suggests that truncation of the TLS protein by translocation may result in the alteration or loss of critical RNA processing capability, leading cells on the path to malignant transformation.

Acknowledgments—We thank Dr. David Ron for TLS cDNA, Dr. Moreau-Gachelin for pCS3-MT-E1A construct, and Dr. Dan Wu for pSG5-Flag vector.

REFERENCES

1. Crozat, A., Aman, P., Mandahl, N., and Ron, D. (1993) *Nature* **363,** 640–644
2. Rabbitts, T. H., Forster, A., Larson, R., and Nathan, P. (1993) *Nat. Genet.* **4,** 175–180
3. Ichikawa, H., Shimizu, K., Hayashi, Y., and Ohki, M. (1994) *Cancer Res.* **54,** 2865–2868
4. Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Houbert, I., DeJong, P., Rouleau, G., Aurias, A., and Thomas, G. (1992) *Nature* **359,** 162–165
5. Bertolotti, A., Melot, T., Acker, J., Vigneron, M., Delattre, O., and Tora, L. (1998) *Mol. Cell. Biol.* **18,** 1489–1497
6. Prasad, D. D., Ouchida, M., Lee, L., Rao, V. N., and Reddy, E. S. (1994) *Oncogene* **9,** 3717–3729
7. Lessnick, S. L., Braun, B. S., Denny, C. T., and May, W. A. (1995) *Oncogene* **10,** 423–431
8. Burd, C. G., and Dreyfuss, G. (1994) *Science* **265,** 615–621
9. Zinszner, H., Sok, J., Immanuel, D., Yin, Y., and Ron, D. (1997) *J. Cell Sci.* **110,** 1741–1750
10. Hallier, M., Lerga, A., Barnache, S., Tavitian, A., and Moreau-Gachelin, F. (1998) *J. Biol. Chem.* **273,** 4838–4842
11. Tsai, S., Bartelmez, S., Sitnicka, E., and Collins, S. (1994) *Genes Dev.* **8,** 2831–2841
12. Vojtek, A. B., Hollenberg, S. M., and Cooper, J. A. (1993) *Cell* **74,** 205–214
13. Caceres, J. F., Stamm, S., Helfman, D. M., and Krainer, A. R. (1994) *Science* **265,** 1706–1709
14. Fu, X. D., and Maniatis, T. (1992) *Science* **256,** 535–538
15. Kozak, M. (1991) *J. Biol. Chem.* **266,** 19867–19870
16. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410
17. Ge, H., Zuo, P., and Manley, J. L. (1991) *Cell* **66,** 373–382
18. Krainer, A. R., Mayeda, A., Kozak, D., and Binns, G. (1991) *Cell* **66,** 383–394
19. Ayane, M., Preuss, U., Kohler, G., and Nielsen, P. J. (1991) *Nucleic Acids Res.* **19,** 1273–1278
20. Stephens, C., and Harlow, E. (1987) *EMBO J.* **6,** 2027–2035
21. Wang, J., and Manley, J. L. (1995) *RNA* **1,** 335–346
22. Hackl, W., and Luhrmann, R. (1996) *J. Mol. Biol.* **264,** 843–851
23. Zinszner, H., Albalat, R., and Ron, D. (1994) *Genes Dev.* **8,** 2513–2526
24. Zhang, D., Paley, A. J., and Childs, G. (1998) *J. Biol. Chem.* **273,** 18086–18091
25. Valcarcel, J., and Green, M. R. (1996) *Trends Biochem. Sci.* **21,** 296–301
26. Zahler, A. M., and Roth, M. B. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 2642–2646
27. Chandler, S. D., Mayeda, A., Yeakley, J. M., Krainer, A. R., and Fu, X. D. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 3596–3601
28. Sun, Q., Mayeda, A., Hampson, R. K., Krainer, A. R., and Rottman, F. M. (1993) *Genes Dev.* **7,** 2598–2608
29. Amrein, H., Hedley, M. L., and Maniatis, T. (1994) *Cell* **76,** 735–746
30. Ring, H. Z., and Lis, J. T. (1994) *Mol. Cell. Biol.* **14,** 7499–7506
31. Heinrichs, V., and Baker, B. S. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 115–120
32. Screaton, G. R., Caceres, J. F., Mayeda, A., Bell, M. V., Plebanski, M., Jackson, D. G., Bell, J. I., and Krainer, A. R. (1995) *EMBO J.* **14,** 4336–4349
33. Jumaa, H., Guenet, J. L., and Nielsen, P. J. (1997) *Mol. Cell. Biol.* **17,** 3116–3124
34. Rossi, F., Labourier, E., Forne, T., Divita, G., Derancourt, J., Riou, J. F., Antonie, E., Cathala, G., Brunel, C., and Tazi, J. (1996) *Nature* **381,** 80–82
35. Neugebauer, K. M., and Roth, M. B. (1997) *Genes Dev.* **11,** 3279–3285
36. McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S. D., Wickens, M., and Bentley, D. (1997) *Nature* **385,** 357–361
37. Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R. V., Gentile, C., Gebara, M., and Corden, J. L. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 6975–6980
38. Mortillaro, M. J., Blencowe, B. J., Wei, X., Nakayasu, H., Du, L., Warren, S. L., Sharp, P. A., and Berezney, R. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 8253–8257
39. Kim, E., Du, L., and Warren, S. L. (1997) *J. Cell Biol.* **136,** 19–28
40. Cote, G. J., Huang, E. S.-C., Jin, W., and Morrison, R. S. (1997) *J. Biol. Chem.* **272,** 1054–1060

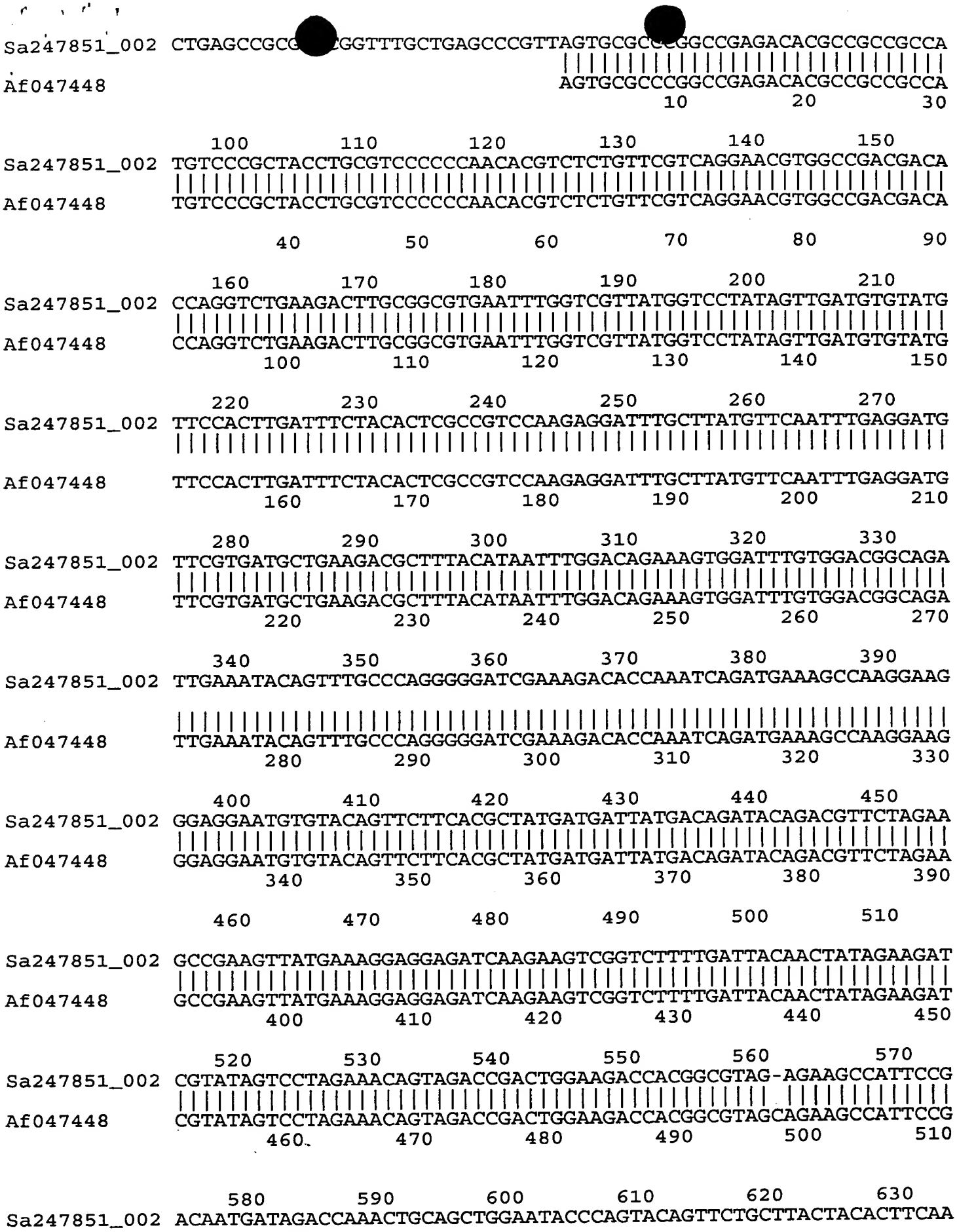
Sa247851_0026.Dna
Emhum1:Af047448

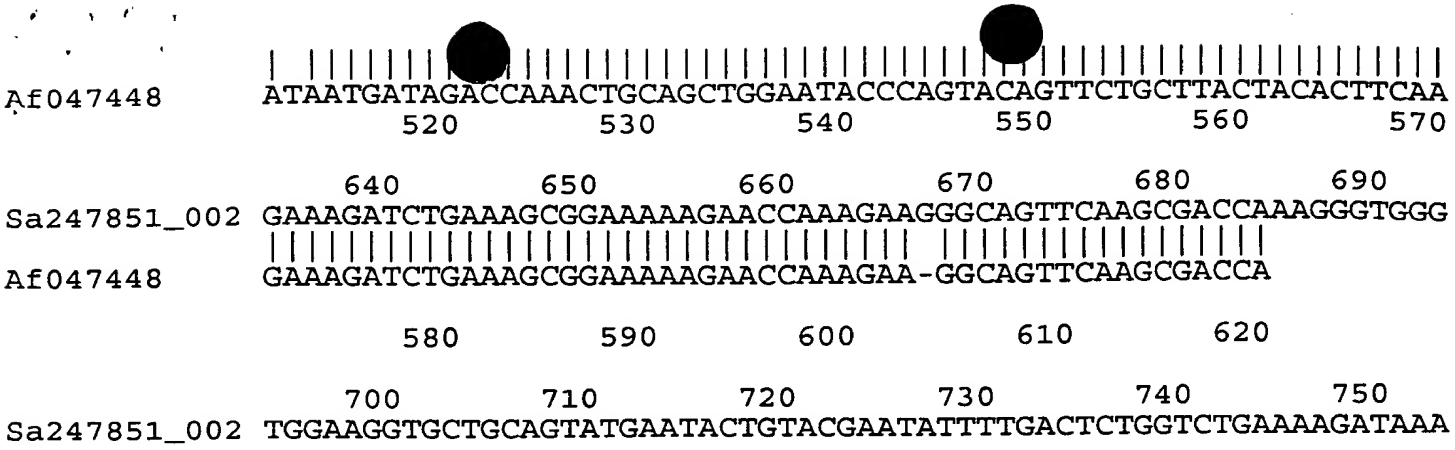
XP2128498

```
ID   AF047448   standard; RNA; HUM; 620 BP.
AC   AF047448;
SV   AF047448.1
DT   17-MAR-1998 (Rel. 55, Created)
DT   05-NOV-1998 (Rel. 57, Last updated, Version 2)
DE   Homo sapiens TLS-associated protein TASR mRNA, complete cds.
KW   .
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;
OC   Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   1-620
RX   MEDLINE; 98447613.
RA   Yang L., Embree L.J., Tsai S., Hickstein D.D.;
RT   "Oncoprotein TLS interacts with serine-arginine proteins involved in RNA
RT   splicing";
RL   J. Biol. Chem. 273(43):27761-27764(1998).
RN   [2]

RP   1-620
RA   Yang L., Embree L., Tsai S., Hickstein D.D.;
RT   ;
RL   Submitted (10-FEB-1998) to the EMBL/GenBank/DDBJ databases.
RL   Medicine/Oncology, University of Washington, 1660 S. Columbian Way, GMR
RL   151, Seattle, WA 98108, USA
DR   SPTREMBL; O60572; O60572.
FH   Key             Location/Qualifiers
FH
FT   source          1. .620
FT                   /db_xref="taxon:9606"
FT                   /organism="Homo sapiens"
FT                   /cell_line="K562"
FT                   /tissue_type="leukemia"
FT   CDS             30. .581
FT                   /codon_start=1
FT                   /db_xref="SPTREMBL:O60572"

FT                   /note="has RNP motif and arginine/serine-rich (RS) domain'
FT                   /product="TLS-associated protein TASR"
FT                   /protein_id="AAC70918.1"
FT                   /translation="MSRYLRPPNTSLFVRNVADDTRSEDLRREFGRYGPIVDVYVPLI

FT                   YTRRPRGFAYVQFEDVRDAEDALHNLDRKWICGRQIEIQFAQGDRKTPNQMKAKEGRN

FT                   YSSSRYDDYDRYRRSRSRSYERRRSRSRSFDYNYRRSYSPRNSRPTGRPRRSRSHSDN

FT                   RPNCSWNTQYSSAYYTSRKI"
SQ   Sequence 620 BP; 185 A; 135 C; 158 G; 142 T; 0 other;
```

SCORES Init1: 2480 Initn: 2991 Opt: 3054 z-score: 1064.8 E(): 0
99.5% identity in 621 bp overlap

40 50 60 70 80 90


Sa247851_002 CTGAGCCGCG GGTTTGCTGAGCCCGTTAGTGCGC GGCCGAGACACGCCGCCGCCA
Af047448 AGTGCGCCCGGCCGAGACACGCCGCCGCCA
10 20 30
100 110 120 130 140 150
Sa247851_002 TGTCCCGCTACCTGCGTCCCCCCAACACGTCTCTGTTCGTCAGGAACGTGGCCGACGACA
Af047448 TGTCCCGCTACCTGCGTCCCCCCAACACGTCTCTGTTCGTCAGGAACGTGGCCGACGACA
40 50 60 70 80 90
160 170 180 190 200 210
Sa247851_002 CCAGGTCTGAAGACTTGCGGCGTGAATTTGGTCGTTATGGTCCTATAGTTGATGTGTATG
Af047448 CCAGGTCTGAAGACTTGCGGCGTGAATTTGGTCGTTATGGTCCTATAGTTGATGTGTATG
100 110 120 130 140 150
220 230 240 250 260 270
Sa247851_002 TTCCACTTGATTTCTACACTCGCCGTCCAAGAGGATTTGCTTATGTTCAATTTGAGGATG
Af047448 TTCCACTTGATTTCTACACTCGCCGTCCAAGAGGATTTGCTTATGTTCAATTTGAGGATG
160 170 180 190 200 210
280 290 300 310 320 330
Sa247851_002 TTCGTGATGCTGAAGACGCTTTACATAATTTGGACAGAAAGTGGATTTGTGGACGGCAGA
Af047448 TTCGTGATGCTGAAGACGCTTTACATAATTTGGACAGAAAGTGGATTTGTGGACGGCAGA
220 230 240 250 260 270
340 350 360 370 380 390
Sa247851_002 TTGAAATACAGTTTGCCCAGGGGGATCGAAAGACACCAAATCAGATGAAAGCCAAGGAAG
Af047448 TTGAAATACAGTTTGCCCAGGGGGATCGAAAGACACCAAATCAGATGAAAGCCAAGGAAG
280 290 300 310 320 330
400 410 420 430 440 450
Sa247851_002 GGAGGAATGTGTACAGTTCTTCACGCTATGATGATTATGACAGATACAGACGTTCTAGAA
Af047448 GGAGGAATGTGTACAGTTCTTCACGCTATGATGATTATGACAGATACAGACGTTCTAGAA
340 350 360 370 380 390
460 470 480 490 500 510
Sa247851_002 GCCGAAGTTATGAAAGGAGGAGATCAAGAAGTCGGTCTTTTGATTACAACTATAGAAGAT
Af047448 GCCGAAGTTATGAAAGGAGGAGATCAAGAAGTCGGTCTTTTGATTACAACTATAGAAGAT
400 410 420 430 440 450
520 530 540 550 560 570
Sa247851_002 CGTATAGTCCTAGAAACAGTAGACCGACTGGAAGACCACGGCGTAG-AGAAGCCATTCCG
Af047448 CGTATAGTCCTAGAAACAGTAGACCGACTGGAAGACCACGGCGTAGCAGAAGCCATTCCG
460 470 480 490 500 510
580 590 600 610 620 630
Sa247851_002 ACAATGATAGACCAAACTGCAGCTGGAATACCCAGTACAGTTCTGCTTACTACACTTCAA


Af047448 ATAATGATAGACCAAACTGCAGCTGGAATACCCAGTACAGTTCTGCTTACTACACTTCAA
520 530 540 550 560 570
640 650 660 670 680 690
Sa247851_002 GAAAGATCTGAAAGCGGAAAAAGAACCAAAGAAGGGCAGTTCAAGCGACCAAAGGGTGGG
Af047448 GAAAGATCTGAAAGCGGAAAAAGAACCAAAGAA-GGCAGTTCAAGCGACCA
580 590 600 610 620
700 710 720 730 740 750
Sa247851_002 TGGAAGGTGCTGCAGTATGAATACTGTACGAATATTTTGACTCTGGTCTGAAAAGATAAA

Sa247851_0001.Prot
Emhum1:Af047448

```
ID   AF047448   standard; RNA; HUM; 620 BP.
AC   AF047448;
SV   AF047448.1
DT   17-MAR-1998 (Rel. 55, Created)
DT   05-NOV-1998 (Rel. 57, Last updated, Version 2)
DE   Homo sapiens TLS-associated protein TASR mRNA, complete cds.

KW   .
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;
OC   Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   1-620
RX   MEDLINE; 98447613.
RA   Yang L., Embree L.J., Tsai S., Hickstein D.D.;
RT   "Oncoprotein TLS interacts with serine-arginine proteins involved in RNA
RT   splicing";
RL   J. Biol. Chem. 273(43):27761-27764(1998).
RN   [2]
RP   1-620
RA   Yang L., Embree L., Tsai S., Hickstein D.D.;
RT   ;
RL   Submitted (10-FEB-1998) to the EMBL/GenBank/DDBJ databases.
RL   Medicine/Oncology, University of Washington, 1660 S. Columbian Way, GMR
RL   151, Seattle, WA 98108, USA

DR   SPTREMBL; O60572; O60572.
FH   Key             Location/Qualifiers
FH
FT   source          1. .620
FT                   /db_xref="taxon:9606"
FT                   /organism="Homo sapiens"
FT                   /cell_line="K562"
FT                   /tissue_type="leukemia"
FT   CDS             30. .581
FT                   /codon_start=1
FT                   /db_xref="SPTREMBL:O60572"
FT                   /note="has RNP motif and arginine/serine-rich (RS) domain"
FT                   /product="TLS-associated protein TASR"
FT                   /protein_id="AAC70918.1"
FT                   /translation="MSRYLRPPNTSLFVRNVADDTRSEDLRREFGRYGPIVDVYVPLD

FT                   YTRRPRGFAYVQFEDVRDAEDALHNLDRKWICGRQIEIQFAQGDRKTPNQMKAKEGRN

FT                   YSSSRYDDYDRYRRSRSRSYERRRSRSRSFDYNYRRSYSPRNSRPTGRPRRSRSHSDN

FT                   RPNCSWNTQYSSAYYTSRKI"
SQ   Sequence 620 BP; 185 A; 135 C; 158 G; 142 T; 0 other;
```

```
SCORES  (3) Init1:  1077 Initn:  1077 Opt:  1077 z-score: 749.5 E():       0
  81.7% identity in 197 aa overlap

                              10        20        30        40        50
Sa247851_000            MSRYLRPPNTSLFVRNVADDTRSEDLRREFGRYGPIVDVYVPLDFYTRRPR
                        ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Af047448      CARPRHAAAL YLRPPNTSLFVRNVADDTRSEDLRRLFGRYGPIVDVYVPLDFYTRRPR
                       10        20        30        40        50        60

                      60        70        80        90       100       110
Sa247851_000 GFAYVQFEDVRDAEDALHNLDRKWICGRQIEIQFAQGDRKTPNQMKAKEGRNVYSSSRYD
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Af047448      GFAYVQFEDVRDAEDALHNLDRKWICGRQIEIQFAQGDRKTPNQMKAKEGRNVYSSSRYD
                      70        80        90       100       110       120

                     120       130       140       150       160       170
Sa247851_000 DYDRYRRSRSRSYERRRSRSRSFDYNYRRSYSPRNSRPTGRPRRREAIPTMIDQTAAGIP
             |||||||||||||||||||||||||||||||||||||||||||||||||| ::       : :

Af047448      DYDRYRRSRSRSYERRRSRSRSFDYNYRRSYSPRNSRPTGRPRRSRSHSDNDRPNCSWNT
                      130       140       150       160       170       180

                     180       190       200       210
Sa247851_000 STVLLTTLQERSESGKRTKEGQFKRPKGGWKVLQYEYCTNILTLV
             :       :::   : |:::: |||||
Af047448      QYSSAYYTSRKIXKRKKNQRRQFKRP
                      190       200
```

Sa247851_0001.Prot
Trembl:O60572

```
ID   O60572      PRELIMINARY;      PRT;   183 AA.

AC   O60572;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TLS-ASSOCIATED PROTEIN TASR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEUKEMIA;
RA   YANG L., EMBREE L., TSAI S., HICKSTEIN D.D.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF047448; AAC70918.1; -.
DR   HSSP; P09651; 1UP1.
DR   PFAM; PF00076; rrm; 1.
SQ   SEQUENCE   183 AA;  22222 MW;  070A7D15 CRC32;


SCORES      Init1:  1071  Initn:  1071  Opt:  1071 z-score: 2509.0 E():      0
Smith-Waterman score: 1071;   100.0% identity in 155 aa overlap

                      10        20        30        40        50        60
Sa247851_000 MSRYLRPPNTSLFVRNVADDTRSEDLRREFGRYGPIVDVYVPLDFYTRRPRGFAYVQFED
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
O60572        MSRYLRPPNTSLFVRNVADDTRSEDLRREFGRYGPIVDVYVPLDFYTRRPRGFAYVQFED
                      10        20        30        40        50        60

                      70        80        90       100       110       120
Sa247851_000 VRDAEDALHNLDRKWICGRQIEIQFAQGDRKTPNQMKAKEGRNVYSSSRYDDYDRYRRSR

             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
O60572        VRDAEDALHNLDRKWICGRQIEIQFAQGDRKTPNQMKAKEGRNVYSSSRYDDYDRYRRSR
                      70        80        90       100       110       120

                     130       140       150       160       170       180
Sa247851_000 SRSYERRRSRSRSFDYNYRRSYSPRNSRPTGRPRRREAIPTMIDQTAAGIPSTVLLTTLQ
             ||||||||||||||||||||||||||||||||
O60572        SRSYERRRSRSRSFDYNYRRSYSPRNSRPTGRPRRSRSHSDNDRPNCSWNTQYSSAYYTS
                     130       140       150       160       170       180

                     190       200       210

Sa247851_000 ERSESGKRTKEGQFKRPKGGWKVLQYEYCTNILTLV

O60572        RKI
```